92-06910/ Final Report
84- 4010
ON05864

URIF/IBM Contract
Final Report

CSILE Project, Centre for Applied Cognitive Science
Ontario Institute for Studies in Education
252 Bloor Street West, Toronto, Ontario M5S 1V6

Project name:
Computer Supported Intentional Learning Environments (CSILE)

Principal Investigators:
Marlene Scardamalia and Carl Bereiter

September, 1990

1. Project Summary and Background

CSILE is a knowledge processing environment which also supports knowledge generation vital for learning with meaning and hence, understanding. At the heart of CSILE is a student-generated database. CSILE supports knowledge manipulation and knowledge construction by helping students work with information in the database, helping them notice gaps, formulate questions, and plan external searches for information for entry into the database.

Educational activities that require knowledge manipulation and generation can have a powerful impact on learning. Fifth and sixth-graders with 6 to 18 months of CSILE experience demonstrate more sophisticated involvement in knowledge-building than is characteristic of children their age. Results suggest that positive learning outcomes will be even greater if children use CSILE at an early age and continue for a greater number of years. Over time, we expect to show improvements in the quality of written work and significant increases in reading comprehension, written expression, and in the maturity and sophistication of students' work in science, history, etc.

CSILE captures students' everyday efforts across the curriculum and contains a wealth of data for analyzing students' contributions to their peers and their involvement in knowledge building. Process measures can be used to identify uses of CSILE that are associated with achievement. CSILE can provide feedback to students and teachers about the numbers and kinds of such activities engaged in by each student. CSILE can also provide feedback and guidance while learning efforts are in progress, not just at their conclusion.

2. Work Under The URIF Contract

2.1 Proposal

Our proposal stated:

> "To realize its full potential, CSILE must be able to function as an environment under which other educational applications can run. The software running under it might include ... compact disk encyclopedias and other reference sources. CSILE's function would be to add a layer of reflective knowledge-building activity on to the intellectual activities supported by these applications. There is evidence that such a ``metacognitive umbrella'' is needed even in the case of sophisticated tutoring systems, to prevent them from being used merely as games."
>
> "With URIF funding we would be able to enhance CSILE to provide interfacing with two existing software components-- a compact disk encyclopedia and a science microworld. Interfacing with an encyclopedia will permit us to experiment with a variety of basic ways of upgrading students' use of reference sources, so that extension to other information resources will be straightforward. ... Procedures for interfacing with the CD encyclopedia and with the selected microworld would be developed with an eye to maximizing generalizability. We expect, therefore, that an interface developed to provide an environment for these two applications could readily be adapted to incorporate other reference sources, other computer microworlds, and various games and instructional programs."

We see this work in the larger context of integration issues in modern computer technology and knowledge work. Integration is turning out to be an important theme in the 1990's, not only because it is vitally necessary for an organization to be able to integrate its disparate information resources in a meaningful and coherent way, but because it is so hard to do. The explosive growth of information technology has created a Tower of Babel of standards and it is often quite difficult to simply textually combine information that follows different standards.

Once those technical challenges of combining information have been met, however, there remains the problem of ensuring, or at least encouraging, the *meaningful* combination of information to produce genuinely useful and valuable knowledge that can be acted or further built on. It is of no use, for example, to combine a table of Canadian rainfall data with a table of Canadian crop production data if the combination is not made with attention paid to the meaningful connections between the two data sets and if appropriate conclusions are not drawn. All too often one finds lists of facts thrown

together on the basis of rather vague associations, with no meaningful connections drawn between them.

2.2 Work Undertaken

We set out to integrate a wide variety of educational resources into the CSILE System in an effective, efficient, and educationally sound manner. Resources considered included:

- text and graphics
- numerical data
- simulations/models

See Appendix A for details of particular information resources considered and Appendix B for details on supporting hardware and software components.

We did this research in the context of hardware and software provided by our corporate partner in this work, the Advanced Workstation Division of the International Business Machines Corporation (see Appendix C).

Over the course of this work, we identified a number of thorny research issues. We have moved towards solutions for several of them and incorporated our solutions into the design for a new overall architecture for the CSILE system, called the Knowledge Building Architecture (see Appendix D).

The research issues that arose were:

- Integration of Information Hosted on Different Platforms
- Integration of Information Represented with Different Data Standards
- Integration of Canadian Material
- Preserving the Source of Imported Information
- Encouraging Coherent Importation of Information
- Discouraging Plagiarism
- Exploiting the Organizing Potential of Reference Information

3. Research Issues Arising From Integrating Reference Material Into CSILE

3.1 Integration of Information Hosted on Different Platforms

The CSILE System is a collection of programs written in C to run under the UNIX™ operating system with the X11 Window System providing graphics and user input support. Unfortunately, almost all educational software is written to run on either the Apple II computer or on the IBM PC with CGA, EGA, or VGA graphics. Without special arrangements, the Apple and IBM educational software cannot run on the same platform as CSILE, making

integration problematic. We did not know whether or not we could find an arrangement that would allow integration effective enough to be viable.

Initially, we hoped to be able to use an IBM PS/2 model 70 with the AIX operating system to run CSILE in one X11 window and to run IBM PC educational software in another X11 window (via the facilities of an AIX program called DOS Merge), letting the student freely move between the two windows, making notes in CSILE about their activities. We anticipated a parallel scheme for the RT PC workstations.

Due to incompatible graphic formats, we were not able to display the output from IBM PC programs in an X11 window, so this scheme had to be abandoned (although we hope to revive it with the RS/6000 workstations, which have an improved version of AIX). It is possible, however, to use a special key combination supported by DOS Merge to toggle back and forth between an X11 Window screen mode and an IBM PC screen mode, meaning that either CSILE or an IBM PC program could be visible on the screen at one time (but not both) and that it was easy and quick to switch between the two.

Forcing users to remember something about their IBM PC program while they switched to CSILE, hiding the PC program, was judged to impose too harsh a memory burden, so an MS-DOS "TSR" (terminate-and-stay-resident) routine was written to allow students to make notes about their use of a program using a "pop-up" editing window that only minimally obscures their view of the PC program's display. The notes written with the editor can easily be imported into the CSILE database.

This scheme works well for interactive programs and reference works delivered on floppy disk, provided that they do not use the Microsoft Windows™ windowing system. For programs that use that system, we created a parallel arrangement with a simple Windows-based text editor, with Windows providing multi-window support.

However, this scheme did not work at all for reference materials delivered via CD-ROM. The MS-DOS software required to drive the CD-ROM reader is incompatible with the AIX DOS Merge facility and so we had to find an alternative arrangement for accommodating different hardware/software platforms.

We devised a looser coupling between the two environments, which will be tested in the classroom during this school year. We equipped the IBM PS/2 model 70 running MS-DOS with Network File System (NFS) software which allows it to seamlessly share files with an IBM RT PC or RS/6000 running AIX, the X Window System, and CSILE. NFS is a widespread standard created by SUN Microsystems and implemented by many manufacturers, so it offers the possibility of integration with a wide variety of systems. With it, a

student can work at the PS/2 "resource workstation" and easily save files produced during use of the available CD-ROM and other reference materials. The shared file system automatically makes these saved files visible to CSILE when the student returns to an RT PC or RS/6000 CSILE workstation.

Due to the popularity of the MS-DOS operating system, and its resultant status as a de facto standard, this scheme has worked well to date, although it has a lower degree of integration than we would like. It works with all reference materials we have assembled, including material input via scanner and optical character recognition software, as detailed in Appendix B.

Three tighter integration schemes are possible, though they are all potentially very expensive:
• construct a more powerful DOS Merge facility that could display IBM PC graphics in an X11 window and that is compatible with CD-ROM drivers. This may not be technically possible.
• construct an MS-DOS version of CSILE that could either "pop-up" for use with other programs or share screen space via a windowing system such as Microsoft Windows™. This may not be technically possible due to memory limitations under MS-DOS.
• port specific programs of interest to the X11 Window System/ UNIX™ platform, where they could run in a window next to a window running CSILE. This may not be feasible except for small programs or for programs for which source code is available. Access to the source code could potentially allow for a very high degree of integration with CSILE, but source code for interesting programs is rarely available. More promising may be custom rapid development tools that would allow us to construct microworlds automatically integrated with CSILE.

Clearly, there remains much fertile ground for research on this issue.

3.2 Integration of Information Represented with Different Data Standards

The graphical images stored by the CSILE database are represented in a particular bitmap format. The graphical images output by the MS-DOS reference sources are in different formats. One of the reference sources, PC Globe, outputs images in a different format and moreover the images are of different dimensions. In order for images from PC Globe to be imported into the CSILE database, they have to be converted into the CSILE graphics format in a manner seamless enough that it would not interfere with the learning activities of CSILE users. Moreover, when it is necessary for the conversion process to result in the loss of information (e.g. when an image has to be clipped), we need to ensure that it is the lowest-value information that is lost.

We have been successful in using a combination of commercially available (HiJaak Plus), free (the UNIX™ fuzzy bitmap manipulation package) and

custom software written in C to handle conversions from graphic formats, especially PCX, used by our MS-DOS reference sources to the graphics format used by CSILE. We have devised procedures to handle the scaling and cropping problems caused by converting images of certain sizes to different sizes. They are based on use of custom programs and a Windows™-based graphics editor. Their effectiveness with student users will be gauged during this school year (using continuing funding from IBM).

During this work, we were struck by the tremendous variety of graphic formats used by MS-DOS software and by CD-ROM's. Some offer specific advantages, such as reduced file size or support for a larger number of colours, but these advantages could be covered by a much smaller set of standards. This situation contrasts sharply with that of the Apple Macintosh, for which there is less than half a dozen graphic formats. It is clear that there are simply too many possible choices under MS-DOS and that this does not serve to create more useful choice for the user, just confusing choice. We will address this issue again later in the report.

3.3 Integration of Canadian Material

At the beginning of this work, we conducted an extensive survey of available CD-ROM's, looking particularly for material that students could use to get a broad view of the world from a Canadian perspective. At this relatively early stage in the development of the medium, there are a variety of very special purpose (and expensive) discs (e.g. containing financial information) but few general-purpose works. We were particularly dismayed by the lack of discs that reflect a Canadian focus -- there are several general-purpose American works, but no Canadian ones. Unfortunately, the availability of Statistics Canada's first CD-ROM, which would be a boon to students, came after the end of the period of the URIF contract. We may be able to find other funding to investigate it, however.

The relative lack of Canadian material on CD-ROM prompted us to investigate alternate means of providing such material to students. For example, the Canadian World Almanac and Book of Facts, produced annually by Global Press in Toronto, contains a wealth of information from many perspectives, especially Canadian, but is unfortunately not available on CD-ROM. We have discussed the possibility of providing this information on-line to students with the publisher and while they do not have any means of assisting us, they have agreed to let us store the text of the book on-line, provided that it cannot be converted back to hard-copy form. While there are likely many copyright holders who would not be as cooperative, there is a wealth of public domain (especially government, e.g. Statistics Canada) information that is only available in print and that would be a powerful on-line resource for students. Consequently, with the help of our corporate partner who has supplied scanner technology, we have investigated optical

character recognition software and are enhancing CSILE to enforce the necessary printing restrictions to bring this resource to the classroom in a way that respects copyright issues.

More generally, it is clear that the costs and complications of CD-ROM production have currently made it very difficult for minority viewpoints to be disseminated via CD-ROM. We see the scanner and OCR software as very important components of a strategy to allow students ready access to alternate viewpoints.

3.4 Preserving the Source of Imported Information

If a student clips a piece of text from a large CD-ROM encyclopaedia and copies it into something he or she is writing in the CSILE database and then another student reads that text and wishes to comment on it, how can the second student know the context of the copied piece of text, so, e.g. they can comment on the appropriateness of the reference? This is made particularly problematic by the nature of facilities provided by CD-ROM retrieval programs, which do not store reference information along with exported material. It is not clear that a mechanism to create this reference information accurately can be put in place without imposing excessive burdens on the student users (e.g. to accurately enter the reference manually).

This issue is the subject of ongoing research. We have arrived at a tentative design. This design is based on recording the keystrokes that the student uses to invoke the retrieval software and to select the items of information for use with CSILE. A simple expert system, augmented for each CD-ROM disc, could analyze the set of keystrokes and produce a reference to be attached to the material to be imported into CSILE. References for material imported from interactive programs could be generated in a similar manner.

Generating references for material imported via optical scanning technology may be more problematic. A tentative design in this area requires the student to scan a book title page or an article title to satisfy a reference-generating expert system (capable of recognizing text that could be used as a reference) that a suitable reference has been provided. An escape mechanism, which allows the student to enter a reference that is later checked by a human, would be needed to accommodate any gaps in the expert system's knowledge.

3.5 Encouraging Coherent Importation of Information

Example: Students are studying the causes of nuclear arms proliferation and means of controlling it. They have browsed through several on-line reference sources and have found a list of countries with nuclear arms and when they developed them and a list of countries with accredited engineering schools. They notice that all nuclear powers have accredited engineering

schools and are tempted to conclude that foreign aid should not be given to non-nuclear powers for the purposes of establishing engineering schools. They overlook the large number of non-nuclear countries with engineering schools. The CSILE design approach raises the possibilities of (a) employing social mechanisms to point out the oversight and (b) argument structure analysis tools to encourage the student to carefully consider cause and effect and to identify which variables in a situation depend on other variables. Automated analysis of arguments is a major research challenge in its own right. We are in contact with researchers at the University of Pittsburgh who are working on this problem.

The overall goal of creating coherent collections of database notes is one of the main goals of the creation of version two of CSILE, currently under way. As a result of the URIF work, we have incorporated special provisions for handling references into the preliminary design for CSILE 2.0. Appendix D is an early design document for this work.

3.6 Discouraging Plagiarism

A commonly identified misuse of on-line reference systems is the "copy-delete" strategy [Brown, Day, and Jones, 1983]. Students browse through an on-line encyclopaedia and copy text and graphics thought to be relevant to their needs. They then delete parts of the copied material, leaving a collage of copied material that they may not fully understand and that neither builds on the original material nor credits the original sources. Existing hypermedia systems, educational or otherwise, do not incorporate mechanisms to discourage this behaviour. Their design is not straightforward and they may be burdensome to the user and so reduce use of on-line reference material.

We have designed a simple mechanism for discouraging plagiarism. Copied material, whether imported from reference sources or copied from other users, will appear within CSILE in a different font or colour. By making it obvious which material has been copied without analysis, and by providing statistical summaries of the amounts of copied material to students and teachers, we hope to leverage social mechanisms to provide a disincentive for plagiarism. No funding is specifically allocated for this research at this time, however.

3.7 Exploiting the Organizing Potential of Reference Information

Example: A student finds a map of the oil-producing countries of the world and wants to use it to organize a database of information about those countries. A mechanism is needed to allow the student to not simply include the map in a collection of material about oil, but to use it to effectively index other information about oil. The user interface for this mechanism has to be

sufficiently general to adapt to a wide range of reference material and yet have a demonstrably positive effect on the indexing of material.

We approached this issue in the context of one of CSILE's ongoing and central problems: providing effective indexing structures to allow users to profitably cluster information. CSILE uses a keyword-based scheme to organize the notes in the shared database. We were interested in determining whether or not imported reference information could aid students in selecting appropriate indexing terms, i.e., keywords, for their notes.

We undertook a controlled study concerning this issue in the spring of 1990. Due to the delays in obtaining IBM equipment for a classroom site by this time, the study took place at one of our other sites which had the requisite equipment. We made specific enhancements to the CSILE System for the purposes of the study.

Briefly, a content expert in the sample domain (Human Evolution) selected a collection of particularly evocative high-quality colour images from material in the domain, along with a set of approximately 70 keywords chosen to be good organizers for the domain. The keywords were positioned on the images, turning the images into *frames* that organized the keywords into a coherent indexing scheme. Students could easily flip between frames and, using a mouse pointing device, click on a keyword to get a definition of the term, to add that keyword to a note they were writing, or to jump to a frame related to that keyword.

We created a human-readable frame description language so that the frame descriptions (pictures, keywords, keyword coordinates, definitions, and keyword-frame relationships) could be authored by a non-programmer.

We divided the test classroom into two groups: control and experimental. Only the experimental group could use the organizing frames to assign keywords to their notes. The CSILE System was modified to prevent the two groups from seeing each other's notes, which would contaminate the results. Extensive records of the students' use of the interface were kept, with every keystroke relating to its use recorded. The experiment lasted five weeks.

The results were very favourable. Students in the experimental group used the organizing frames to assign keywords more appropriate to their notes and to assign more keywords per note. Encouraged by these results, we are planning more general use of reference material as organizing frames.

4. Conclusions

URIF funding contributed to two large areas of CSILE activity:
- the reference information integration activities described above
- the design of a new version of CSILE, CSILE 2.0, the basis of which is described in Appendix D.

In this section, we cover conclusions in the former area and leave most discussion of the latter to Appendix D.

4.1 Standards Widely Accepted by the Marketplace Are Important

The issues discussed in sections 3.1 (Integration of Information Hosted on Different Platforms) and 3.2 (Integration of Information Represented with Different Data Standards) point to a need for more widespread standards for data. Specifically, we need standards for enhanced text (e.g. with font information), graphics, and indexed data (e.g. on CD-ROM). The last would avoid the need to have a different retrieval program for each CD-ROM database and, also importantly, would allow retrieval programs running on different hosts to access the same CD-ROM's. We have found that the lack of standards can lead to a great deal of effort bridging different standards or, in some cases, can halt progress completely.

Governmental agencies at all levels could help immensely in this effort. Governments hold huge amounts of data which could be very useful if it were available in a usable form on CD-ROM. By adopting a consistent and small set of standards for this data, government could take the lead and create more order out of the current chaos, particularly in the MS-DOS world.

4.2 The Importing of Reference Material Holds Promise in Several Areas

We have found that the promise of effective integration of reference material lies in two areas:
- a means of including attributed facts in documents produced by knowledge workers. Replacing guesses or poorly remembered facts with verifiable facts should improve the quality of knowledge produced considerably. For this to work, it is important to be able to follow the reference link back to its original source.
- a means of organizing facts or other material through the use of organizing frames, such as maps, timelines, diagrams, and so on. Often, the key to reaching effective conclusions is finding a good organizational frame. One can envisage a CD-ROM full of a collection of indexed organizational frames, suitable for organizing a wide variety of information.

4.3 Measurement of the Effectiveness of Integration Schemes Should Include a Cognitively-Based Evaluation of the Coherence of the Integrated Information

The CSILE design philosophy stresses that information should be *coherently* combined, not just slapped together. Indeed, if the information is not combined with meaning, the result may be worse than if the information were not combined at all, since a lack of knowledge is often better than misleading knowledge.

The ongoing efforts to create CSILE 2.0 (see Appendix D) are aimed at producing software tools to help knowledge workers create coherently structured work. The approach taken is not to rely on artificial intelligence techniques to automatically and coherently integrate information, but to amplify the social mechanisms used by experts to produce coherence. We do this with the *aid* of a computer technology (likely including artificial intelligence) and the ultimate goal of creating not just coherent collections of knowledge but of users of CSILE better able to deal with knowledge in general.

4.4 The Economic Impact of Information Integration Research Could Be Dramatic

We contend that effective information integration is the key to useful knowledge, which is the key to vastly improved productivity. In many of our discussions and dealings with our corporate sponsors, we are struck by the depth of their concern about the difficulty of coherently integrating information and doubly struck by how urgently their large bodies of knowledge, distributed over different parts of their organizations, need to be coherently integrated.

We expect that the spinoffs from this work will put Ontario in a highly advantageous position in this field. Blessed with a singularly progressive educational system, which stresses the cooperative learning style necessary for modern knowledge work, Ontario is in an excellent position to take advantage of developments in networking technology and the kinds of development reported here to create radical improvements in productivity.

5. References

Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. Child Development, 54, 968-979.

6. Appendices

Appendix A: Information Resources

We looked at a wide variety of material suitable for importing into CSILE:
- electronic data (text and/or graphics), delivered via floppy disks or CD-ROM
- data produced by interactive programs (some of which use CD-ROM)
- print-based material (books, newspapers, magazines, etc.)

We surveyed each of these areas.

A.1 Electronic text and graphics data

We conducted an extensive survey of available CD-ROM's. Despite the lack of discs that reflect a Canadian focus, we were able to find some discs that are relatively neutral in their outlook or useful enough that teachers would feel their bias is outweighed by their usefulness. To date, we have investigated a summary collection of statistics about different countries (compiled by the US Central Intelligence Agency), the electronic form of the Grolier encyclopaedia, the Oxford English Dictionary, several modern dictionaries (including one specific to science and technology), a database of food characteristics, and a summary of news events of the 1980's.

We have also explored reference material delivered via other electronic media. In particular we have investigated a floppy-disk package called PC Globe, which contains a world map, many national and international statistics, and effective means for exporting the information to other programs such as CSILE. We are also exploring the possibility of obtaining Statistics Canada data on magnetic tape and storing it on hard disk for student access. We feel it is important to expose students to a wide variety of information, as their working lives will be spent dealing with many different kinds and sources of information.

A.2 Data from Interactive Programs

The information presented by an interactive program, e.g. a simulation of a nuclear reactor, can constitute valuable reference material. Some educational software programs are called "microworlds", indicating that they are a simulation of a small part of the world. This can be a relatively concrete part of the physical world (e.g. a nuclear reactor, a spring) or a more abstract part of the world (e.g. part of mathematics).

We surveyed a number of microworlds, with an emphasis on microworlds that address mathematics education. We examined the programs Mathematica, Calculus, The Calculus Toolkit, Cactusplot, Graphing Equations, Milo, The Geometric Supposer, and LogoWriter and a few other

programs such as spreadsheets (Excel), drawing packages (Corel Draw, Microsoft Paintbrush), animation packages (AutoCad Animator), and a computer-assisted design package (Generic CADD 3D). Many observations came out of this work in three areas: educational suitability, software compatibility and level of integration. Many microworlds are not educationally suitable since their user interfaces are either intended for university students and are thus too complex (e.g. Mathematica) or for very young students and thus offer little educational challenge.

A.3 Print-Based Material

We have investigated the use of print in conjunction with a high-quality optical scanner (the Hewlett-Packard ScanJet Plus) and an advanced optical character recognition package (OmniPage 386 version 2.1) running on the IBM PS/2 workstation. The combination of these two products has the ability to accurately enter a wide variety of text, from books, newspapers, magazines, and brochures. If the type is large enough (approximately 10 pt or more), then acceptably low error rates are encountered. In addition, the OmniPage interface is simple enough to use that we expect it to be successful in the classroom environment, as it has already been successful in our adult knowledge-worker environment.

A.4 Importing Processed CSILE Notes From Other Programs and Exporting

If CSILE is truly to be an umbrella which increases the educational effectiveness of other software, then in addition to allowing material to be imported into its student database, it must also support the export of material from the student database so that it can be used with programs from third parties. A good example is a sophisticated spelling, grammar and style checker that we have investigated, RightWriter. A student may want to export a piece of text written within CSILE to be analyzed by RightWriter. The text will then be revised, within CSILE, on the basis of the critique prepared by RightWriter.

The desire to create highly polished collections of material from the CSILE database, e.g., as "projects", also motivates exporting. CSILE is not designed to facilitate presentation-oriented aspects of knowledge manipulation, so tasks normally associated with desktop publishing or presentations must be done with other software. An effective exporting mechanism for CSILE would allow a student to mark a number of CSILE notes as material to be exported for use in, e.g., a student newspaper, and then make it easy to use a page layout program such as Microsoft Word for Windows to assemble the material into, e.g., a visually pleasing newspaper format.

We have created procedures to allow students to export text and graphics from CSILE in such a way that they can be easily handled by MS-DOS

programs. The effectiveness of these procedures will be evaluated with students this fall.

Appendix B: Resource Workstation Configuration

The IBM PS/2 Model 70 workstation (25 MHz 80386 CPU, 8Mb RAM, 120Mb hard disk) supplied by IBM Canada for use in this research has been augmented with hardware and software to support information integration.

Hardware added:

- NEC Multisync 3 colour monitor - to support CGA, EGA, and VGA colour graphics
- NEC CDR-77 CD-ROM drive and interface card - to read CD-ROM's
- Hewlett-Packard ScanJet Plus optical scanner and interface card - for integration of print material
- Ungermann-Bass NIC Ethernet card - to provide connectivity between the resource workstation and CSILE workstations

Software added:

Basic system software
- IBM DOS Version 3.30
- Disk Manager DMDRVR.BIN hard disk driver for D:
- NEC SCSI Interface kit CD-ROM software version 1.12 with MSCDEX.EXE Microsoft's DOS operating system extensions for CD-ROM drives
- PC-NFS (Network File System) Version 3.0.1 - network connectivity
- Microsoft Windows Version 3.0 - graphical user interface environment
- Hijaak Plus - graphics conversion software

Development software
- Turbo C++ compiler
- resident_C - to implement "TSR" programs
- Turbo C Tools - general purpose library of routines for C

CD-ROM reference materials
- Microsoft Bookshelf
- The CIA World Factbook
- The Oxford English Dictionary
- McGraw-Hill Concise Encyclopedia of Science and Technology and the Dictionary of Scientific and Technical Terms
- The New Grolier Electronic Encyclopedia
- News Digest / Facts On File

Software for importing other reference materials
- PC Globe 3.0 - maps and data on all countries

• OmniPage/386 Version 2.1 - optical character recognition software

Software for processing material exported from CSILE, for later importing
• RightWriter Version 3.1E - grammar and style checker
• Microsoft Excel - spreadsheet package for data analysis

Software for preparing polished work using material exported from CSILE
• COREL Draw Version 1.2 - sophisticated drawing package
• Generic 3D Drafting Version 1.1 - three-dimensional computer-aided design package
• MathCAD Version 2.5 - symbolic mathematics tool
• Autodesk Animator - animation package

Appendix C: Collaboration with International Business Machines Corp.

Our work with IBM (Advanced Workstation Division) during the URIF contract period (January 1989 - July 1990) involved:
• the design of appropriate workstation configurations for classroom installation of the CSILE program on advanced IBM workstations
• the porting of CSILE to the workstations
• preparation for the installation of CSILE facilities in the school site
• research into the creation of artificial intelligence techniques to assist students in navigating through the CSILE database

During the period of this research, our contract coordinator at IBM's Advanced Workstation Division has been very excited by our work and has put a great deal of effort into ensuring that the hardware platform supplied by IBM for this work (a) will have a long life and (b) is powerful enough for the research. Accordingly, once she discovered that the RT PC workstation, initially selected as the base for the work, would be discontinued by IBM, she looked for alternatives, including the PS/2 model 70 running AIX (IBM's version of UNIX™) and the AIX-based successor to the RT workstation, the very high-performance RS/6000. Three of our staff attended a confidential briefing on the new machines in July 1989. We agreed with our coordinator that these machines would significantly advance our work. Originally, IBM envisaged a mixed PS/2 and RS/6000 strategy but they have now opted for school and development lab installations consisting almost totally of RS/6000 machines.

As interim development equipment, in 1989 IBM provided three IBM PS/2 model 70-A21's (16Mb of memory, 120Mb disk storage, Ethernet cards, high-resolution display cards and 19" colour monitors), an IBM laser printer, and, in 1990, two heavily loaded RT model 135 workstations. Due to possible delays in obtaining stable system software for the RS/6000 machines, we have opted for an RT PC based installation for the fall of 1990 at our IBM school site, the University of Toronto Schools high-school. The nine RT PC

workstations for this site have arrived, their installation is nearing completion, and the process of educating teachers in their use has begun.

We have also recently received some RS/6000 workstations for development purposes and they amply live up to their reputation for high performance. We are very excited by the prospects for a very high-performance school installation of these machines in 1991, using continued funding provided by IBM.

Our thrust in the area of artificial intelligence is predicated on successive refinement of strategies that use the impressive computational power of the IBM equipment to analyze the student database. With the recent arrival of the new equipment, we now look forward to a period of experimentation with such strategies. (Despite the conclusion of URIF funding, IBM funding continues, so we will be able to make continued progress on the issues covered in this report, albeit at a very much reduced rate due to the need to allocate resources to other aspects of our work with IBM).

Appendix D: A Knowledge-Building Architecture for Computer-Supported Learning[1]

Marlene Scardamalia and Carl Bereiter

Introduction

CSILE (Computer-Supported Intentional Learning Environments) is intended to help students actively construct and develop their knowledge of the world (Scardamalia & Bereiter, in press a; Scardamalia, Bereiter, McLean, Swallow, & Woodruff, 1989). This purpose requires that we define an architecture for knowledge building that, by definition, does not presume the capabilities it wishes to foster.

In its simplest form, CSILE is a networked system that gives students simultaneous access to a database that is composed of text and graphical notes that the students produce themselves and means of searching and commenting on one another's contributions. Our development strategy has been to get a simple version of the system into full operation as quickly as possible and then to investigate experimentally the kinds of enhancements that best support the objective of cooperative knowledge construction. Students and teachers have been working with us for five years in this initial experimental phase. A number of enhancements have been tried, some retained, some dropped, others put on hold because one cannot experiment with too many innovations in the same year without disrupting classroom education.

Throughout this initial development period, experimentation has dealt with additions to or minor changes in the fairly simple initial architecture, which we call CSILE 1.0. On the basis of what has been learned during this period, we are now designing a much more complex architecture known as KBA, "knowledge-building architecture," and which will be embodied in CSILE 2.0.

In this document we briefly survey the strengths and weaknesses of core features of CSILE 1.0 and then devote the bulk of the document to describing and explaining the design ideas that we aim to incorporate into CSILE 2.0.

---

[1]Paper presented at the NATO Advanced Research Workshop, "Computer-Based Learning Environments and Problem-Solving," Leuven, Belgium, September, 1990.

1 Strengths and Weaknesses of CSILE 1.0

1.1 Evaluation of Principal Features

Table 1 summarizes major features of CSILE 1.0 and lists observed strengths and weaknesses of each.

The centerpiece of CSILE 1.0 is its student-generated, communal database, and all of our experience to date encourages us to retain it as such in CSILE 2.0. With no modifications whatever, the system has proved educationally usable and useful for students from grade one through graduate school. A grade 1 student may produce a simple graphical note, which may look like a scribble to other people but which the child wants to preserve. Thus comes the challenge of storing the note, so a label is needed--perhaps just a letter or two. Once stored, simple search routines enable students to retrieve the note. They do not even have to recall their label, they simply must recognize it from a list of entries. This same list gives them an overview of what others are doing, and through simply placing a cursor on a label they can retrieve their own notes or view notes of others. They can then browse the database and even copy into their personal space parts of notes generated by others and use these as building-blocks for subsequent notes that they write. Each step is relatively easy and rewarding. We have found that the notion of labeling a note, having it disappear, then being able to retrieve it is not simply fun, it is a real boon to early literacy. Students come to realize that it is easier to find a note if they label it effectively; further, others are more likely to find it helpful. Notes are stored automatically with the author's name, and can be retrieved by searching author fields. Accordingly, even the youngest of students soon learn to read the names of their peers. Through these simple mechanisms they are brought into contact with many notes more sophisticated than they themselves can produce, or possibly even read. They get ideas for future notes and it is our impression that they purposefully browse with the intent of finding notes more sophisticated than they produce. In any event, the student-generated communal database provides them with many models of work across a wide spectrum of ability levels.

Given these basic components, students in the early elementary years have worked together to produce linked charts and to write plays. The very same environment has grade six students entering and searching notes to work out differences in breathing apparatuses for different animals, leading to a collection of interrelated notes. Graduate students have used the same environment to enter notes about course readings and to compare different interpretations of texts. To date, students and teachers have invented countless uses.

At all levels, students come to see themselves as authors of public-domain material. Every note in the database is put there by users; thus they come to think of themselves as contributors to a large body of information, with the attendant responsibility to produce material worthy of entry into a communal medium. They are not simply navigating through notes entered by teachers or professional authors. Results suggest that over time this has a natural upgrading effect. Even students previously identified as learning disabled and immigrants who speak very little English (a significant population in the schools where we work) become prolific users of CSILE. In five years of testing with hundreds of students, grades one through university, we have not found a student unable or unwilling to contribute. Much of the material is also fascinating to read, thus encouraging reading and written commentary. Students become active readers, writers, and illustrators.

The weaknesses of the student-generated databases, as indicated in Table 1, arise from the sheer quantity and diversity of notes that accumulate over a school term. A major challenge for design of CSILE 2.0 is to increase coherence, selectivity, and accessibility, without losing the benefits of a large, undifferentiated database in which "everything goes."

All the other features itemized in Table 1 we also judge to be well worth preserving but improving in CSILE 2.0. Author-assigned keywords do not enjoy high status in information technology generally, and we would agree, if the focus is exclusively on information retrieval. But for educational applications it is equally important to consider what keyword assignment does for the author. We believe it plays an important role in getting students to consider knowledge as an object, which is a key requirement for knowledge workers of all kinds. In addition to assigning keywords, students also assign "thinking types" to their notes, using student-designed icons. Thinking types refer to the kind of knowledge-processing that went into producing the note--for instance, whether it was pursuing a question, creating a plan, developing a new idea, or reporting acquired information. In CSILE 2.0, such simple labeling functions will be elevated to choices of special environments for carrying out particular kinds of knowledge-building processes.

1.2 Overall Evaluation of CSILE 1.0

A variety of formal and informal evaluations of the educational effectiveness of CSILE 1.0 have been carried out, although systematic experimental-control evaluations were only initiated in the past year and have not yet yielded findings. In outline form, this is what evaluations to date indicate:

A. Overall, signs of significant advantage for students in CSILE:

1. Question Asking--students readily pose questions that challenge experts, even about topics that are unfamiliar (not necessarily a result of CSILE per se, but of question-asking activity that CSILE supports).

2. Writing--depth of explanation

3. Learning Goals--recognize where they are likely to have difficulty learning and kinds of goals they might attain.

4. Epistemological terms--increase with time

B. Additional, informal observations

1. Writing--Increase in quality and quantity of text written by students.

2. Illustrations--We have no comparative data, but complexity and frequency of illustrations is impressive. Our belief is that graphics facilities open up a new knowledge-representing avenue. Students spontaneously link illustrations and text.

3. Motivation--stays high for the whole year (for multiple years for those in longitudinal study). They get better the more time they spend.

4. Special populations (students just learning English as their second language and students independently assessed as having learning difficulties) seem to thrive in such an environment.

2 CSILE 2.0: Implementation of an Advanced Knowledge-Building Architecture

2.1 Overview of the Knowledge Building Architecture

CSILE 1.0 may be thought of as providing a single environment in which various operations may be carried out upon knowledge dealing with any of a variety of themes. The major architectural change in CSILE 2.0 is the creation of distinct environments for different kinds of knowledge operations and provision for distinct "thematic spaces" differentiated according to the topic or substantive problem (e.g., dinosaur extinction) that work is focused on.

The concept of a single community database is preserved, however. All notes are accessible from any environment or thematic space, although the ways they are linked and the things that can be done with them vary.

The relation between environments and thematic spaces may be visualized as a grid, as shown in Figure 1, with columns representing environments and rows representing different thematic spaces. A user may move horizontally from environment to environment while remaining in the same thematic space or may remain within a given environment and move vertically through different thematic spaces.

To convey more explicitly how thematic spaces and knowledge-building environments are interrelated, we will develop an example using material from an actual CSILE 1.0 database. One graphic note (Figure 2) represents a student's attempt to show how the diaphragm works. Related graphic and text notes (Figure 3) by the same student are attempts to understand the role of the diaphragm in hiccupping. Another child is trying to figure out what happens inside your body when you sneeze. This child's notes are presented in Figure 4 We also see elaborated notes (Figure 5) which the student characterizes as "new learning" (via a "thinking type" icon). These present information discovered through research about how the diaphragm works and its relationship to hiccupping. Other notes in the database (not shown) deal with the relationship between the diaphragm and laughing. To pursue the full set of ideas in the database, different students need to be brought together to elaborate their diagrams and notes. For example, the role of the epiglottis needs to be elaborated, as does the question of why a different sound is associated with a hiccup than with a sneeze. There seems little doubt that these students would explore these increasingly penetrating issues if they had these notes all conveniently available for joint consideration. That is the idea of a thematic space. It is analogous to a "conference" in the communications world.

The thematic space gets students and notes concerned with a particular issue together, but then what do they do? They may just want to go on adding ideas and information related to the phenomena of hiccupping, sneezing, coughing, and laughing. In that case, they can stay in the HOME environment, which in CSILE 2.0 will correspond to the undifferentiated environment encompassing the whole of CSILE 1.0. But they may want to work more systematically toward a deeper and more explicit account of how the diaphragm creates these varying effects. For that purpose, they can move (while remaining within the same thematic space) to the HOW-IT-WORKS environment. In effect, they will have taken all their notes about the diaphragm with them, but the HOW-IT-WORKS environment will provide a framework for assembling notes and creating new notes that explain a process at increasing levels of specificity. It will also provide tools for producing slide-show sequences, simple animations, and simulations, which they may

find useful in demonstrating their conceptions of the processes. They may develop or read about competing theories, which they wish to compare and evaluate. For that purpose, they can move into an EXPLANATIONS environment, which provides ways of highlighting the explanatory aspects of notes and for investigating the coherence of complex explanations. Thus, thematic spaces and knowledge-building environments have complementary roles, which combine to support the focusing of cooperative knowledge-building effort on problems and issues and the pursuit of knowledge to deeper levels.

In the following sections we provide a more detailed account of KBA, describing its five main components: the community database, knowledge-building environments, thematic spaces, tools and procedural facilitations, and background operations.

2.2 Community Database

A single community database underlies all the thematic spaces and knowledge-building environments (Scardamalia, Bereiter, McLean, Swallow, & Woodruff, 1989).

- All of the knowledge stored in the community database is stored in the form of user-generated notes, although these notes may be created in a number of ways.

- All notes are treated the same, regardless of whether they are text, graphic, copied from some external source, or generated as output of some knowledge-building tool.

- Information is brought into particular environments in the form of notes and the outcomes of knowledge-building operations carried out within environments are preserved in the form of other notes. Thus the database at any time provides a representation of the totality of the public knowledge constructed by the user group.

- Material from other databases must be converted into notes before it can be stored in the community database. This requires that the user add keywords or other kinds of information that constitute "knowledge value-added," so that the note is never simply a cut-and-pasted fragment.

2.3 Knowledge-Building Environments

We aim in the design of these environments to foster the following kinds of advanced knowledge processes: (a) testing the explanatory power of hypotheses; (b) exploring analogies; (c) identifying causal relationships; (d) illustrating and animating objects and thus identifying underlying mechanisms; (e) working out the historical relevance of ideas; (f) analyzing numerical relationships in terms of their relevance to principles of mathematics; (g) extracting domain vocabulary from notes and creating background environments where these terms can be clustered and linked to create an overview of note contents, (h) engaging in research through integrated use of on- and off-line reference material, and so on.

The following are brief functional descriptions of some of the environments envisioned for CSILE 2.0.

2.3.1 HOME environment

The HOME environment is the default environment that the user enters when signing on to CSILE. From the HOME environment, users can branch directly to any other environment or can branch to a thematic space. Being in the HOME environment and not being in any particular thematic space is equivalent to being in the CSILE environment in its current version. We would expect that for young and new-to-CSILE students this may be the only environment available initially, with other environments and thematic spaces available as they gain sophistication. All resources for creating, retrieving, and commenting on notes are available in this environment.

2.3.2 The EXPLANATION environment

The motive for entering this environment should be a desire to understand some aspect of the world or some puzzling facts. The EXPLANATION environment's interface will require that notes created or brought into the environment be given short legends which summarize the fact or hypothesis they discuss. The interface also provides a means of linking these into a network of the kind that ECHO (a PDP program that evaluates networks of propositions according to principles of explanatory coherence) can operate on (Thagard, 1989). This means that links indicating "explains" or "contradicts" can be created or removed. The interface may also provide procedural supports to help users think through the creation of links.

2.3.3 The HOW-IT-WORKS environment

The focus in this environment is on explaining how things work--machines, physiological processes, geological processes, economic, political, etc. Like the EXPLANATION environment, it supports the building of coherent

explanations. However, it is more suited to working out and elaborating explanations than with pitting one explanation against another.

The main challenge in creating this environment is to design an interface that facilitates a query-and-response process that leads to progressively deeper levels of process description. It would seem that a hierarchical zoom-in structure, as in current Charts, would be suitable. A given chart could show, in comic strip fashion, three main steps in a process. Clicking on one of the steps would bring up a chart showing a finer breakdown of that step; clicking on one of those substeps would bring up a finer breakdown still, and so on. Other tools that might be available in the HOW-IT-WORKS environment would permit the construction of flowcharts, simulations, and animations.

2.3.4 The DATA EXPLORATION environment

This is an environment for exploratory analysis of data. It would include graphing tools, perhaps a spreadsheet program, and kinds of data exploration such as are available in Minitab. An important challenge in design of this environment is to maintain close connections between the data and what they refer to, so that users do not get lost in number-crunching.

2.3.5 The MEANINGS environment

The focus in this environment is on domain vocabulary and its interrelationships. From an educational standpoint, this is an important aspect of work within any thematic space.

A central feature of the MEANINGS environment is a concept map that is automatically generated for any thematic space. Suppose the students who developed notes on the diaphragm and its relation to sneezing, hiccupping, etc., had been in CSILE 2.0 and had established a "sneezing-hiccupping-laughing" thematic space. When a student is in that space and moves into the MEANINGS environment, the first thing that would appear would be a map showing those keywords and a number of other words and a network of links connecting one word with another. This map would have been generated and updated from time to time on the following basis: Excluding a specified set of high-frequency words, every pair of words co-occurring in a sentence in any note in the "sneezing-etc." thematic space would appear on the map as two nodes with a line linking them. (Co-occurrence within sentences has been found to be a good basis for identifying the significant vocabulary of a domain.) A central word like "diaphragm" should have many links and it should be possible to follow links from it to most other words on the map. However, there would probably also be many stray pairs or clusters of words unconnected to the main cluster. Some of these stray words might not really belong on the map, having co-occurred fortuitously. Others, however, might have important connections that have not yet been

made explicit in any notes. Certain words and links on the map are highlighted. Clicking on them will bring up a note that defines or explains the word or the link between a pair of words. These notes are not attached automatically but are attached by students working in the MEANINGS environment.

There are four things a student might do relating to the concept map for a thematic space:

1. Tag a word as not belonging on the map. When a certain number of students have tagged a word in this way, it is deleted from the map, along with its links.

2. Tag a link as unimportant or irrelevant. When a certain number of students have tagged a link in this way, it is deleted, but the words connected by it remain, provided they still have links to some other words.

3. Create a link between presently unlinked words by writing a note in which the two words are used within the same sentence.

4. Attach an explanatory note to a word or link. This may be a new note or a note that already exists. Linking requires filling in a slot that indicates what word or link the note explains.

Through these actions, the knowledge map grows in extent and coherence, providing a representation of the collective knowledge of the students who have contributed to the thematic space. The map also provides a basis for live group discussions, in which names for various links can be proposed, discussed, and later actually assigned.

A further level of semantic investigation is foreseen, when CSILE 2.0 is linked to external (especially CD-ROM) information sources. Students may then designate external documents related to their thematic space and have CSILE extract the domain vocabulary that links to words on their own concept map. This additional vocabulary, appearing as shaded nodes and links in the background of their own map, provides clues to further relationships to explore.

2.3.6 The EXPERT environment

An analysis of questions by students in CSILE classrooms shows that approximately 35% (Scardamalia & Bereiter, in press b) of their questions are ones that might require help from an outside expert to answer. The EXPERT environment will aim to connect them with experts in their field of inquiry, while treating experts' time as a limited resource that must be used wisely.

The challenge in creating the interface or educational layer will be to help students filter and upgrade material to be sent to experts. Questions should not be going to experts until users have taken advantage of the resources available to them, and in the course of doing this they should come to appreciate the fact that they can make progress without expert help even on very difficult questions.

The center of the EXPERT environment is, of course, an external communication facility. But knowledge-processing tools for outgoing communication will serve to synthesize and filter requests for information. For instance, when too many questions have accumulated for submitting to a particular external expert, the environment will provide means for students to comment on priorities and suggest revisions and ways to consolidate several questions into one. (Educationally, this pre-consultation work could prove as valuable as the actual consulting with an expert.) Responses from experts need to be converted into notes in order to become part of the community database. According to the "knowledge value-added" principle, this requires that information be accompanied by some elaboration or interpretation by a student rather than just be pasted in as a note. The EXPERT environment will provide convenient means and procedural supports for adding such knowledge value to received information.

2.3.7 The DOCUMENTS environment

This environment serves two main functions: Preparing notes to become candidates for "published" status in the database, and preparing larger documents that are to be printed as reports, library entries, "take-homes", or whatever. Mainly this environment provides a range of editing and publishing tools beyond the basic ones available in all the other environments--e.g., grammar and style checkers, pagemakers, facilities for bringing separate notes into the environment and assembling them into larger documents.

2.4 Thematic Spaces

When students sign on to CSILE they will enter the HOME environment. After completing such HOME environment activities as reading comments and checking their calendars, they have the choice of either going into a thematic space or going to another environment. Suppose they elect to go into a "dinosaur extinction" thematic space. If they then decide to search for notes, the note search screen will inform them that there are, say, 24 notes in the "dinosaur extinction" thematic space and they will have the option of calling up all of them or of carrying out an ordinary search. However, the search screen will only present them with keywords and other identifiers (such as authors) associated with those 24 notes. If they want to search more

widely they will have to choose an option that lets them search "outside the thematic space."

Knowledge Organizers for Thematic Spaces. Knowledge organizers are visual frames, such as maps, timelines, and graphs that appear as the initial screen when one enters a thematic space. The knowledge organizers are labeled with keywords. Clicking on a keyword label permits attaching the keyword to a note in the write-a-note window and getting other information related to the keyword. CSILE needs to provide for these organizers to be created in several ways: (1) teacher assigns a list of keywords, students produce the visual frame and place the keywords in it; (2) the organizer is built from a digitized image taken from book, etc., and students attach keywords; (3) students build the whole organizer from scratch; (4) a previous organizer is used (retrieved from a library of these) and is edited for the current purpose; (5) the organizer is built using a template map, timeline, etc., drawn from a library of such templates.

The template option is especially desirable because it opens up the possibility of subsequently merging knowledge organizers. For instance, one group of students might develop a thematic space about African artists, using a map template to situate the artist's names geographically. Another group might do the same with African writers, political leaders, etc. A merging of these maps, with the different sets of keywords appearing in different colors, would allow for consideration of African nations that appeared to be especially active culturally.

A more futuristic possibility involves knowledge-building tools that could retrieve and compile information on templates at the command of the user. Students could write simple natural-language programs to combine information in informative ways, such as, "Combine maps of Western Europe. Show names of poets in red and names of inventors in green." If students also used timelines when recording poets and inventors they could additionally specify that they'd like to see only those poets and inventors born between specified dates. With yet a bit more intelligence (in CSILE that is), they could do novel searches such as "Find all poets and inventors born in medieval times and combine on map of Western Europe."

2.5 Tools and Procedural Facilitations

Tools as Knowledge-Building Resources

The basic idea of a knowledge-building environment is that knowledge is brought into the environment and something is done to it that enhances its value. Tools available in the environment are means by which this value-adding work is done. Consequently, the design of environments and tools is

guided by the goal of maximizing the value they add to knowledge--either the public knowledge represented in the community database or the private knowledge and skill of the user.

Some tools, such as basic text and graphics tools, are available in all environments. Others, however, are specific to certain knowledge-building environments and are integrated into the particular emphasis of the environment. We have mentioned ECHO as a key tool in the EXPLANATIONS environment, providing a means for students to evaluate the coherence of linkages between explanations and facts. But it is integrated into the environment by means of an interface that gets students to identify and label explanations and "facts needing explaining" as notes are brought into this environment. These labels then become the tokens that ECHO operates on. STELLA (Richmond & Peterson, 1988), a general process-modeling tool, is a natural to integrate into the HOW-IT-WORKS environment; but again, we shall wish to design the environment's interface so that it provides a natural way for the information in notes to be translated into the data structures that STELLA works with.

As these examples illustrate, the architecture of CSILE 2.0 will be able to employ a variety of software tools designed by others. But we aim for CSILE to provide much more than an access route to these tools. CSILE's environments should provide a reason to use the tools, a context in which their use contributes to the overall development of the community's knowledge, and supportive means both for translation of information from notes into input to the tools and translation of output from the tools into notes that become a meaningful part of the community database.

Procedural Facilitations

The interface characteristics we have just described represent an important form of procedural facilitation. Procedural facilitation is a concept we have been promoting for some years and that is often incompletely or poorly understood (Bereiter & Scardamalia, 1987; Scardamalia & Bereiter, 1985). As applied to software, procedural facilitation refers to any design characteristics that enable users to carry out more complex or sophisticated mental operations than they could otherwise carry out. An equation-solver is not a procedural facilitation, because the software, not the user, solves the equations. Thus, the user can get by with less complex mental operations than otherwise. An effective "help" system is usually not in itself a procedural facilitation, because it helps one use the software, not perform the mental operations for which the software provides a medium (e.g., word processor "help" does not help you write better). On the other hand, some people find that a good outliner is a procedural facilitation, because it enables them to organize their thinking better. Good project management software

might have a similar effect, although the software we have seen is so rigid and intrusive that it disrupts thought.

Designing good procedural facilitations, even in such simple forms as hints that pop up on the screen, is not easy. The computer is intervening in complex and delicate cognitive processes, where it is easy to do more harm than good. That is why we have been very cautious about introducing procedural facilitations and why we are moving in CSILE 2.0 toward designing whole environments so that they will be supportive of certain kinds of mental activity, rather than depending on specific prompts or structuring devices that are triggered by actions of the user. There remains a role for specific procedural facilitations, however, and we will illustrate two of them:

1. Procedural Facilitations for Commenting

The basic design of the main CSILE screen has proved to be supportive of commenting. The two-window design permits the student to have a note to be commented on in one window and to compose the comment in the other, thus aiding relevance and coherence. With no further supports, students produce many helpful comments, but these tend to consist of very global evaluations ("good") and low-level suggestions, such as pointing out spelling errors. We have been experimenting with off-line suggestions of sentence-openers that encourage more constructive comments: e.g., "My hypothesis is different from yours. I think ...," "I found this reference that explains ..." These, together with encouragement by teachers, have substantially increased both the quantity and quality of comments. CSILE 2.0 will provide menus of such sentence-openers, which students can select and pop into their comments whenever they wish. These menus will be teacher-authorable, so that the sentence-openers can be adapted to the age-level of students and to local needs.

2. Keyword Supports

As indicated earlier, we see keywords as having educational value both for the author who assigns them and for the students who use them in searching. But assigning relevant and useful keywords is notoriously difficult, therefore a good target for procedural facilitation. CSILE 2.0 will provide automatic support in checking the spelling of new keywords and screening for redundant variations (plurals, etc.). We also want to provide a kind of procedural facilitation, however, that will help students think more effectively about the content of their notes. One simple support is the following: When a student highlights any word in a note and clicks the "keyword help" icon, CSILE searches for that word in the concept maps that have been created in the MEANINGS environment (q.v.). If it finds the word there, it retrieves the portion of the map surrounding that word, thus

displaying the other domain-relevant words that have already been associated with it. This may suggest (a) that the word is already connected to other words relevant to the present note, and so it should be a good keyword to assign; (b) that the words already connected to this word are in quite a different domain (which could mean that the word is a poor choice or that it is a good keyword to assign because it will bring the note to the attention of students working on a different topic); (c) that the word has presently little connection (or no connection if no map is retrieved) to other words in the database, so that if it is used as a keyword it should probably be accompanied by other keywords that have wider currency. A more advanced form of help system could, once it obtains a recognizable highlighted word, search an on-line thesaurus to suggest related keywords, including ones that are more general and ones that are more specific than the highlighted word.

An Expanded Role for Thinking-Type Icons: From Thinking Types to Knowledge Contributions

CSILE is intended to make students more aware of the kinds of knowledge they are dealing with, and there is evidence--from increases in the use of knowledge-related words, for instance--that it has that effect. We expect the number of knowledge concepts with which students actively work to increase dramatically as they deal more with "facts needing explaining," "hypotheses," "data," etc., in CSILE 2.0. Whereas CSILE 1.0 provided icons for students to label their notes according to "thinking types," CSILE 2.0 will provide for labeling "knowledge contributions." This change is in keeping with the focus on public knowledge and its growth. It is not so much what went on in the author's head when composing the note that needs to be recorded but what the note is believed to contribute to the community knowledge base. The list of possibilities could include standard types that are the same in every environment and special types relevant to the particular environments (like "facts needing explaining" in the EXPLANATIONS environment). Whatever ones are chosen, the student is required to complete a brief label, which provides a short statement of the idea, question, fact, or whatever. These entries then become the objects that various tools can work with. There could, for instance, be tools that will display a list of the high-level questions associated with a certain keyword. The screen would show the 30-character-or-so question labels, and the user could then select from this screen any complete notes that looked interesting. Thus the assignment of knowledge types would play a useful role, and mislabeling would show up as notes appearing where they didn't belong.

2.6 Background Operations

Background operations refer to computations that CSILE performs automatically rather than on the command of the user. Thus, for example,

spelling checking is not a background operation when it is called for by the user (instead, we refer to the spelling checker as a tool available from all environments). But if CSILE automatically checks new keywords for spelling, that is a background operation (even though it may use the same program as the spelling-checking tool). Three classes of background operations may be distinguished. They differ in how closely they interact with the user, and consequently they differ in how fast they have to be able to run:

1. Automatic checking. This refers to operations such as checking new keywords for spelling and checking new keywords, propositions, or other distinguishing attributes for redundancy (for example, checking to see whether a new hypothesis in the EXPLANATION environment is semantically similar to one that already exists). These checks are ordinarily run at the time the user is storing a note, and therefore they have to be performed rapidly enough that they do not delay the storing process excessively. In general, automatic checking background operations involve some search and comparison and usually some uncertainty, so that the system must interact with the user before an action is taken.

2. Suggesting and reminding. An important class of background operations does automatic analyses of note attributes or of text content and then informs the user of such things as related notes that may be of interest, domain vocabulary relevant to note content, or uses that other people have made of the user's notes. Although speed is a virtue with these background operations, so that information can be provided to users at a time when it is most relevant, speed is not as important as with automatic checking, because no processes are held up waiting for the operation to be completed.

3. Monitoring. Included in monitoring are all the kinds of data compiling and analysis that are currently being done off-line and used in preparing reports. The expectation is that in the new architecture these kinds of monitoring will be part of CSILE (although perhaps still carried on after hours) and will provide the basis for reports that can be called for by researchers, teachers, and students. A special kind of monitoring that does not exist presently will involve the assigning of activation values to notes. These values will not only be helpful for assessment reports but will also be used in the programs that carry out suggesting and reminding operations.

Through background operations running in CSILE, authors of related notes are notified of the new activity regarding their notes, or authors working on a similar topic are notified of each other's work. Background operations can collect information on individual students that can go into reports to or from teachers. They can also collect information on individual notes that can affect

a note's status. For example, one design we are working on involves something similar to citation indices. Notes gain "credit" as they are used--read, copied, commented on, or attached to a larger structure in some knowledge-building environment. Additionally, if material is copied from a note the copied information will be entered into the new note in a different font, with a bibliographic note recorded automatically. Such easy cross referencing and tabulating of sources should enhance the sense of collaboration in knowledge building.

In designing functions using background operations, everything possible needs to be done to defeat the "Big-Brother-is-watching-you" syndrome. Users need to be provided with an accurate model of the kinds of data used in background operations, so that they know what effect actions on their part can have on them. One might think of the design challenge as that of allowing teachers and students to track the information and kinds of educational activity they believe are valuable. When this is tracked with ease and consistency across users the information can serve to make explicit otherwise little known conditions for improvement.

2.7 Importing Material from On-line Reference Material

In CSILE 2.0, on-line reference material will include CD-ROM, video, microworlds, and a broad array of commercially available software. The central educational problem to be addressed is how to allow students to take good advantage of these resources, and yet avoid mindless cutting and pasting of reference material.

To date, CSILE users have not had access to clip art, video, reference material on CD-ROM or other resources that would allow wholesale importation of material into the student generated database. However, they do take notes from reference books and then spend a good deal of their CSILE time copying their notes into CSILE. This is clearly wasteful of time on CSILE. But on the positive side they do add to the notes in interesting and oftentimes significant ways. First, they tend to add commentary, to elaborate ideas, to make inferences and generally to interpret ideas. That is, there is an interpretive layer to their notes, even though a good deal of the information might be taken directly from reference material. Second, they encode the note, selecting keywords so they can store and retrieve the note easily. Accordingly, there is value added to all notes that are in CSILE.

We want to ensure that future designs preserve the "Value-Added" principle and, as much as possible, ensure that students actually process information they copy into CSILE. The solution will be specially designed import protocols that place a layer of educational activity between imported material and the student-generated database itself. Accordingly, material from external

databases will be required to be converted into notes before it can be stored in the community database. This minimally requires adding a keyword and a label indicating the nature of its knowledge contribution. When material is copied into the student-generated database it will not remain unless there is sufficient "value added." We may want to require addition of some original text before it can be stored, or we may want to use background operations to assign low credit to a copied note, so that it needs to be enhanced or used in significant ways (e.g., needs to acquire "citation credits," as described in the preceding section) in order to avoid dropping out of the public database via a built-in decay factor. None of this is quite sufficient to keep students from copying junk and adding gibberish to it, however, so we need to be attentive in aspects of design to encourage attention to the knowledge value of notes and to discourage the mindless production of collages.

Summary and Conclusions

Although the constructivist view of learning is well on its way to becoming a part of conventional educational wisdom, there are few circumstances, even in the most avant-garde schools, that actually encourage students to view what they are doing as the construction of knowledge. Conscious, cooperative development of shared knowledge is the focus of CSILE. Results to date, both from our own work and from similarly-motivated work (Brown & Campione, in press; Lampert, 1986, 1989) convince us that elementary school students can profitably make knowledge construction the focus of their efforts, although it is a novelty to them (and to their teachers) and requires a great deal of support.

Our development strategy has been to put a simple version of CSILE into full operation as quickly as possible and then to investigate experimentally the kinds of enhancements that best support the objective of cooperative knowledge construction. On the basis of what has been learned we are designing a much more powerful system, a knowledge-building architecture, which will be embodied in a second generation of CSILE. A major architectural change in CSILE 2.0 will be the addition of distinct environments for different kinds of knowledge operations. The basic idea of a knowledge-building environment is that knowledge is brought into the environment and something is done to it that enhances its value. Tools available in the environment are means by which this value-adding work is done. Consequently, the design of environments and tools is guided by the goal of maximizing the value added to knowledge--either the public knowledge represented in the community database or the private knowledge and skill of the user.

CSILE 2.0 will also permit the importing of information from electronic media such as CD-ROM and electronic mail, whereas in the current version

of CSILE everything that enters the community database is authored by the students. We want to pursue the knowledge value-added principle here, as well, so that material is not mindlessly pasted into the CSILE database without the information's having been processed and enhanced in some way by the student.

Knowledge-building environments of the kind proposed have no real precedent in educational practice, even though many of the tools included in the planned environments already exist and have proved useful in differently structured contexts. The environments must do much more than provide access to knowledge-processing tools. A knowledge-building environment should provide a reason to use the tools, a context in which their use contributes to the overall development of the community's knowledge, and supportive means both for translation of information from notes into input to the tools and translation of output from the tools into notes that become a meaningful part of the community database. Finally, and perhaps most importantly, the instructional program as a whole must be designed in such a way that systematic pursuit of problems of meaning, explanation, "how-it-works," etc., are integral to the curriculum.

References

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Campione, J. C. (in press). Communities of learning and thinking or a context by any other name. *Human Development.*

Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction, 3*(4), 305-342.

Lampert, M. (1989). *Teaching and learning division for understanding in school.* East Lansing, MI: Department of Education, University of Michigan, Mimeo.

Richmond, B., & Peterson, S. (1988). *A user's guide to STELLAStack.* Lyme, NH: High Performance Systems, Inc.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Vol. 2. Research and open questions* (pp. 563-577). Hillsdale, NJ: Lawrence Erlbaum Associates.

Scardamalia, M., & Bereiter, C. (in press a). Schools as knowledge-building communities. In S. Strauss (Ed.), *Vol. 5. Human Development.* Norwood, NJ: Ablex.

Scardamalia, M., & Bereiter, C. (in press b). Higher levels of agency for children in knowledge-building: A challenge for the design of new knowledge media. *Journal of the Learning Sciences.*

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. *Journal of Educational Computing Research, 5*(1), 51-68.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*(3), 435-502.

TABLE 1
CSILE 1.0: Features, strengths, and weaknesses

Feature	Strengths	Weaknesses
Student-Generated Database		
- simultaneous access for net-worked users -one database spans curriculum -notes are the unit of knowledge -text and graphic notes	-constructivist approach to learning -appropriate for all ages -pride in ownership -cooperative knowledge building	-too much material for teachers to monitor -redundancy and clutter -early drafts mixed with finished -wide range of quality
Keyword Assignment		
-required to store note -may be selected -may be typed in -readers may add keywords	-aid to early literacy -aid to thinking about knowledge	-idiosyncratic keywords -misspellings -minor variations--plurals, articles, etc. -keywords used to uniquely identify note -long keyword lists
Note Statuses		
-anonymous vs. public -candidate for publication -published (requires review process)	-anonymity permits "trial baloons" -publication encourages quality, significant contributions, teaches revision -readers may restrict search to published notes	-publication generally based on quality judged in isolation rather than in relation to aggregate knowledge -attaining published status not sufficiently motivating
Searching and Reading		
-pseudo-natural language search -boolean search with multiple criteria: keywords, thinking type, note status, author -reader can page through retrieved notes or obtain listing	-usable by all ages -flexible in meeting varying needs and preferences	-overuse of author as search criterion -multiple keywords seldom used -difficulty of retrieving single note encourages assignment of idiosyncratic keywords
Commenting		
-links note in Write-a-note window to note in Read-a-note window -new comments retrieved when author signs on	-powerful means of cooperative knowledge-building -benefits of mail system without the chit-chat	-commenting not available on graphic notes -supports needed to upgrade value of comments
Linked Charts		
-labels attached to one graphic activate links to others -permit hierarchical zoom-in /zoom-out -other users may produce charts to complete links	-encourages elaborated knowledge structures -good medium for cooperative work -draws attention to knowledge relationships	-similar function not available for text notes -restricted to tree structure

Figure 1

Environments and Thematic Spaces in KBA

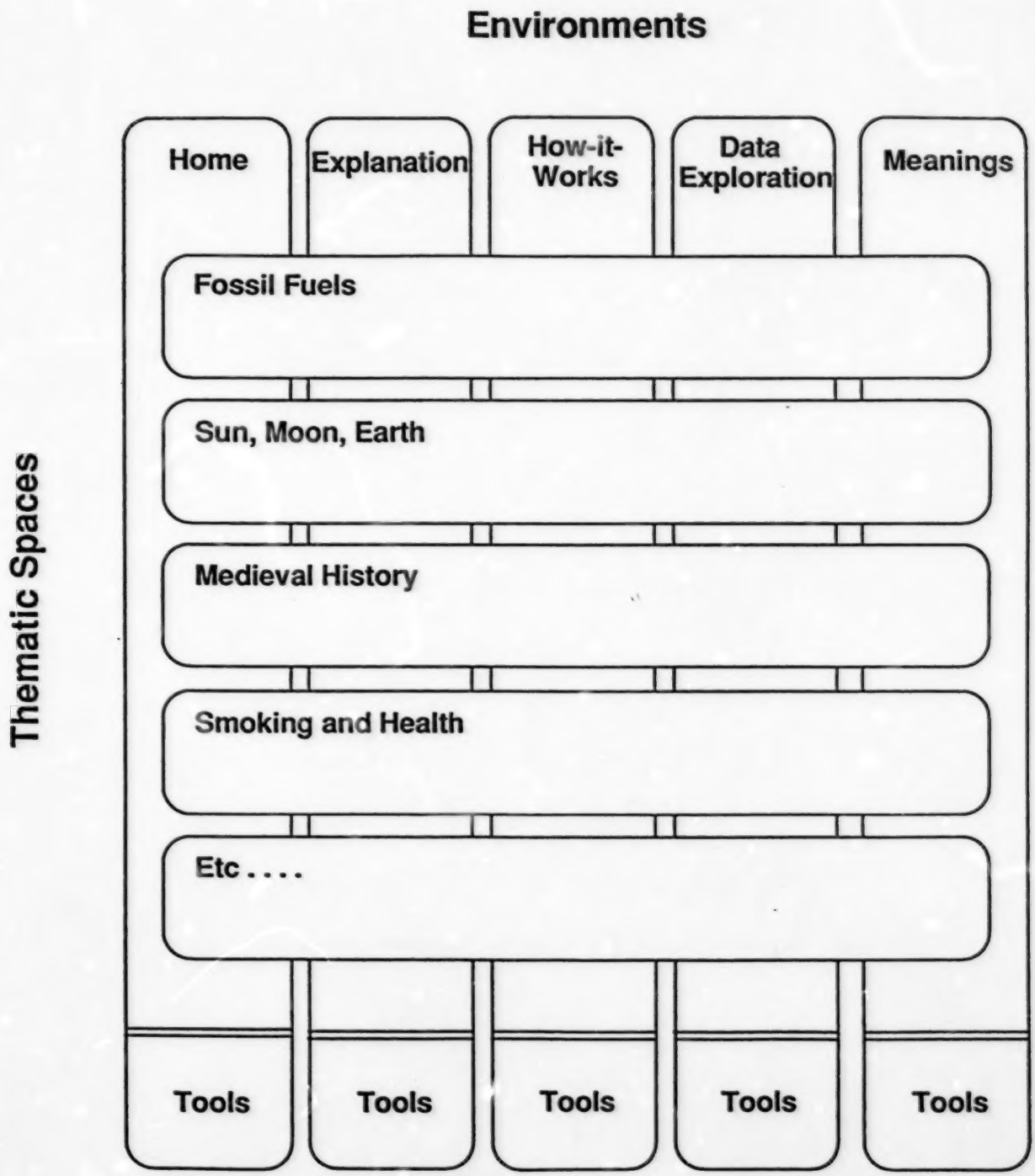
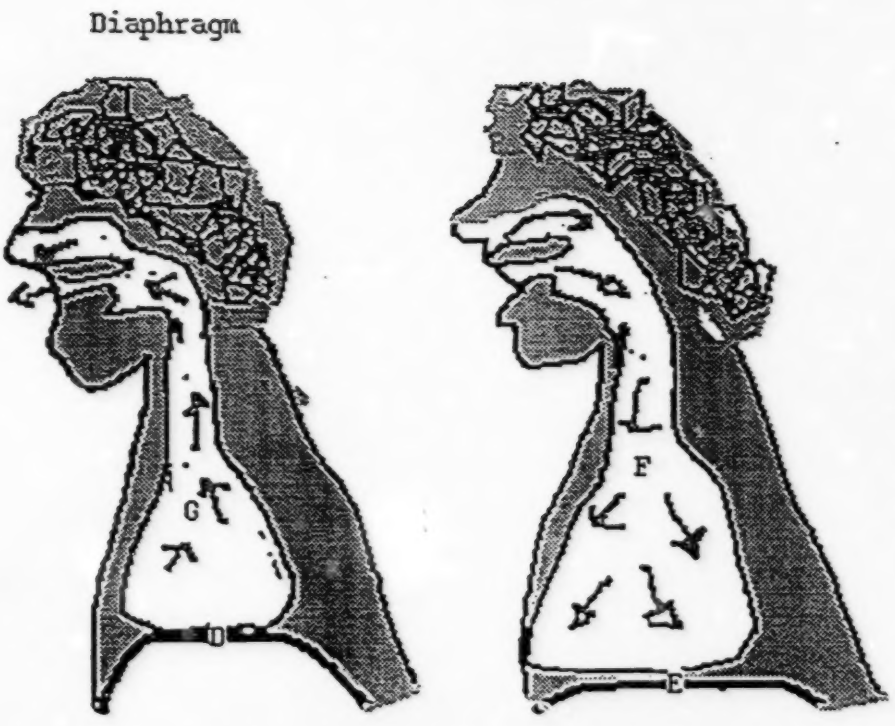


Figure 2

+a: Anatomaical Post
+b: hiccuping
+c: laughing
D: Diaphragm exhale
E: Diaphragm inhale
F: inhale
G: exhale

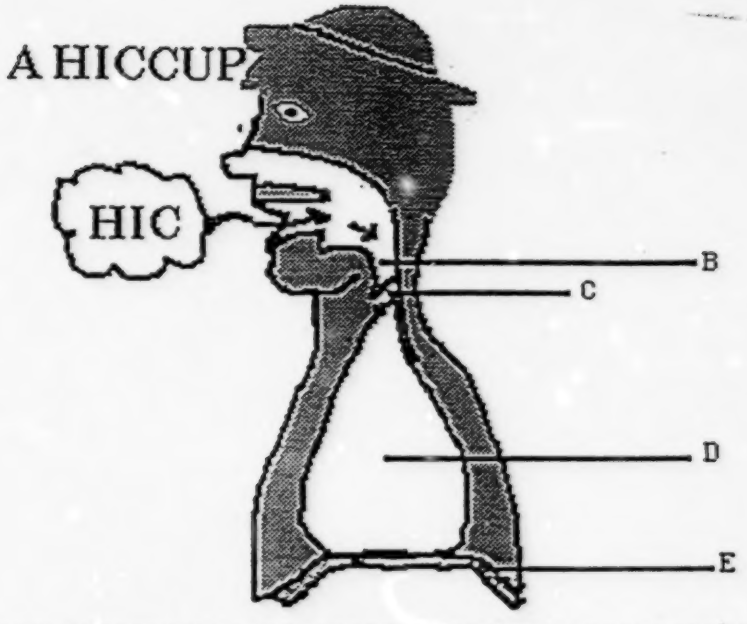


This is a picture of how the diaphragm works.

Figure 3

+a: Diaphragm
B: TRACHEA
C: EPIGLOTTIS
D: LUNG
E: DIAPHRAGM

hiccuping



Author:
Thinking Type: problem
Note Level: SHARED
Topic: Biology
Keywords: BIOLOGY, diaphragm, hiccuping, laughing, problems.

PROBLEM

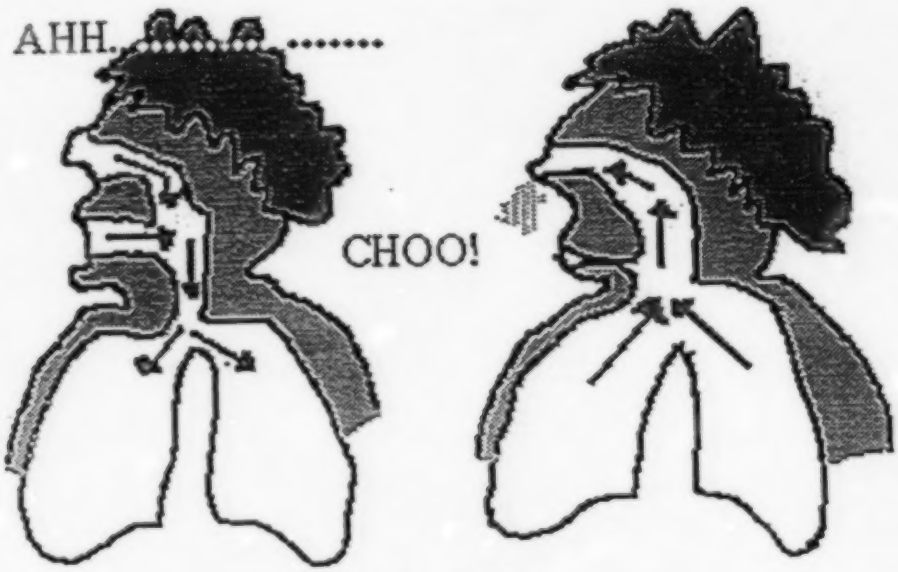
How is the diaphragm connected to hiccuping and laughing?

Why do we hiccup?

Figure 4

+a: Anatomical posit

SNEEZES



SNEEZING

By

see text
by

Author:
Thinking Type: high-level question
Note Level: CANDIDATE
Topic: Biology
Keywords: sneeze, BIOLOGY

Sneezing

Problems: What is a sneeze?

Questions:

1. What is happening inside your body when you sneeze?
2. Do sneezes have a history?
3. What is "snot"?

Figure 5

Author:
Thinking Type: new learning
Note Level: CANDIDATE
Topic: Biology
Keywords: BIOLOGY, Muscles, diaphragm, pacer, paralyzed.

The Diaphragm

The diaphragm is a sheet like muscle made up of string-like muscles. It is located right under your rib-cage and without it you would not be able to breath. The diaphragm goes up and down when you breath:when you inhale the diaphragm tightens and flatens and when you exhale the diaphragm loosens and curves upwards.

If you were to paralyze part of your diaphragm which can happen if the diaphragm nerve gets severed (usually during surgery) it might be o .k.. because there is something that is called a pacer which they can surgically attach to the nerve which is serving the diaphragm scince the nerve is what causes the paralysis in the diaphragm by obviously malfunctioning and not sending a message to the diaphragm the pacer would take its place. But if somehow you were to paralyze your whole diaphragm (which is possible by the way) then you would be in trouble for that is how some drugs cause death!

Author:
Thinking Type: new learning
Note Level: SHARED
Topic: Biology
Keywords: BIOLOGY, diaphragm, hiccups.

HICCUPS

As you know from reading my note on the diaphragm the diaphragm has a certain rhythm. A hiccup is caused by a sudden spasm of the diaphragm. At the same time in a fraction of a second a small flap at the back of your throat called your epiglottis closes blocking off the trachea and when the air tries to to get in it bounces of the epiglottis and makes a "HIC" sound.